Customer No. 000959

A

Case Docket No. AGM-002

THE ASSISTANT COMMISSIONER FOR PATENTS
Box Patent Application
Washington, D.C. 20231

"Express Mail" Mailing Label Number EL 008 489 077 US

Date of Deposit January 25, 1999

I hereby certify that this transmittal letter and the papers referred to as being enclosed therein are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231.

Viriato G. Cardoso

Signature

Viriato G. Cardoso

Please Print Name of Person Signing

Sir:

Transmitted herewith for filing is the patent application of

Inventor(s): Frank Kastenholz

For: NETWORK PACKET FORWARDING LOOKUP WITH A REDUCED NUMBER OF MEMORY ACCESSES

Enclosed are:

☐ This is a request for filing a ☐ continuation ☐ divisional application under 37 CFR 1.53(b), of pending prior application serial no. ___________ filed on ______________ entitled ___________________.

☒ 17 pages of specification, 5 pages of claims, 1 pages of abstract.

☒ 23 (Figures 1-20) sheets of drawings.

☒ An Executed Declaration, Petition and Power of Attorney.

☒ An Executed assignment of the invention to Argon Networks, Inc. (1p.). A recordation form cover sheet (Form PTO 1595) is also enclosed.

☒ A verified statement to establish small entity status under 37 C.F.R. 1.9 and 37 C.F.R. 1.27.

☐ Other________________________________

The filing fee has been calculated as shown below:

FOR:	NO. FILED (Col. 1)	NO. EXTRA (Col. 2)
BASIC FEE	///////////////	///////////////
TOTAL CLAIMS	31 - 20	= 11
INDEP. CLAIMS	8 - 3	= 5
☐ MULTIPLE DEPENDENT CLAIMS PRESENTED		

* If the difference in Col. 2 is less than zero, enter "0" in Col. 2.

SMALL ENTITY RATE	SMALL ENTITY FEE		OTHER THAN SMALL ENTITY RATE	OTHER THAN SMALL ENTITY FEE
/////////	$ 380	OR	////////	$
x 9=	$ 99	OR	x 18=	$
x 39	$ 195	OR	x 78	$
+130	$	OR	+260	$
TOTAL	$674.00	OR	TOTAL	0

☐ Please charge my Deposit Account No. 12-0080 in the amount of $.
A duplicate copy of this sheet is enclosed.

☒ A check in the amount of $ $674.00 to cover the filing fee is enclosed.

☒ The Commissioner is hereby authorized to charge payment of the following fees associated with this communication or credit any overpayment to Deposit Account No. 12-0080.
A duplicate copy of this sheet is enclosed.

- [x] Any additional filing fees required under 37 C.F.R. 1.16.
- [x] Any patent application processing fees under 37 C.F.R. 1.17.

- [ ] The Commissioner is hereby authorized to charge payment of the following fees during the pendency of this application or credit any overpayment to Deposit Account No. 12-0080.
A duplicate copy of this sheet is enclosed.

- [ ] Any patent application processing fees under 37 C.F.R. 1.17.
- [ ] The issue fee set in 37 C.F.R. 1.18 at or before mailing of the Notice of Allowance, pursuant to 37 C.F.R 1.311(b).
- [ ] Any filing fees under 37 C.F.R. 1.16 for presentation of extra claims.

- [x] A check in the amount of $ 40.00 to cover the recording of assignment documents is also enclosed.

- [x] Address all future communications (May only be completed by applicant, or attorney or agent of record) to Kevin J. Canning, Esq. at **Customer Number: 000959** whose address is:

Lahive & Cockfield, LLP
28 State Street
Boston, Massachusetts 02109

Date: January 25, 1999

LAHIVE & COCKFIELD, LLP
Attorneys at Law

By Kevin J. Canning
Kevin J. Canning, Esq.
Reg. No. 35,470
28 State Street
Boston, MA 02109
(617) 227-7400
Telecopier (617) 742-4214

09/237259.012599

Applicant or Patentee: Frank Kastenholz
Serial or Patent No.: ______
Filed or Issued: Filed Herewith
Title: NETWORK PACKET FORWARDING LOOKUP WITH A REDUCED NUMBER OF MEMORY ACCESSES

Attorney's Docket No.: AGM-002

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) and 1.27(c)) - SMALL BUSINESS CONCERN

I hereby declare that I am

☐ the owner of the small business concern identified below:
☒ an official of the small business concern empowered to act on behalf of the concern identified below:

NAME OF SMALL BUSINESS CONCERN Argon Networks, Inc.
ADDRESS OF SMALL BUSINESS CONCERN 25 Porter Road
Littleton, Massachusetts 01460

I hereby declare that the above identified small business concern qualifies as a small business concern as defined in 13 CFR 121.12, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees to the United States Patent and Trademark Office, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention described in:

☒ the specification filed herewith with title as listed above.
☐ the application identified above.
☐ the patent identified above.

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights in the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d), or a nonprofit organization under 37 CFR 1.9(e).

*NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

NAME ______
ADDRESS ______
☐ INDIVIDUAL ☐ SMALL BUSINESS CONCERN ☐ NONPROFIT ORGANIZATION

NAME ______
ADDRESS ______
☐ INDIVIDUAL ☐ SMALL BUSINESS CONCERN ☐ NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statement were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING Steve Willis

TITLE OF PERSON OTHER THAN OWNER Chief Technical Officer

ADDRESS OF PERSON SIGNING 25 Porter Road, Littleton, Massachusetts 01460

SIGNATURE Steve Willis DATE 1/20/99

NETWORK PACKET FORWARDING LOOKUP WITH A REDUCED NUMBER OF MEMORY ACCESSES

Related Applications

This application claims the benefit of priority under 35 U.S.C. 119(e) to co-pending U.S. provisional application Serial No. 60/090,028, filed June 19, 1998, the entire contents of which are hereby incorporated by reference.

Technical Field

The present invention relates generally to switches and routers and more particularly to a network packet forwarding lookup with a reduced number of memory accesses.

Background of the Invention

Computer networks have typically been viewed as being divisible into several layers. The Open Systems Interconnection (OSI) reference model established by the International Standards Organization (ISO) defines a computer network as having seven layers. Figure 1 depicts the seven layers that are defined by the OSI reference model. Layer one is the physical layer, which is responsible for transmitting unstructured bits of information across a link. Layer two is the data link layer. The data link layer is responsible for transmitting chunks of information across a link. Layer three is the network layer. The network layer is responsible for enabling any pair of systems in the computer network to communicate with each other. Layer four is the transport layer. The transport layer is responsible for establishing a reliable communication stream between a pair of systems. Layer five is the session layer, which is responsible for offering services such as dialogue control and chaining. Layer six is the presentation layer, which provides a means by which applications can agree on representations for data. Layer seven is the application layer in which applications such as file transfer services and management services operate.

The Internet protocol (IP) is a layer three network protocol. The IP protocol is a messenger protocol that is part of the Transmission Control Protocol (TCP)/IP protocol suite. TCP is transport layer protocol that facilitate reliable byte stream communication. IP sets forth an addressing scheme that is useful in tracking Internet addresses for different nodes, recognizing incoming messages and forwarding outgoing messages. Each IP packet is a data packet that contains header information and a payload.

IP addresses are 32 bit globally unique addresses that are generally represented in a dotted decimal notation where the dots (i.e. periods) separate the four bytes of the address. An example of an IP address in dotted decimal notation is "1.2.3.4." Although an IP address is a single 32 bit value, each IP address contains two pieces of information. As shown in Figure 2, each IP address 10 contains a network identifier 12 and a host identifier 14. The host identifier identifies the host system to which the IP address is assigned. The network identifier identifies the network in which the host system resides.

In order to appreciate how IP addresses are used, it is helpful to consider an example. Figure 3 shows an example of a computer network in which IP packets are sent between host 20 and host 24. In this example, host 20 is part of network 1 and host 24 is part of network 2. A number of switching nodes 22 interconnect network 1 with network 2. These switching nodes may be switches and/or routers that forward IP packets between network 1 and network 2. Host 24 is host number 97 within network 2. Thus, expressing the address of host 24 in <network, host> form, the IP address for host 24 is <2, 97>. IP packets are forwarded from their source to their destination on a hop by hop basis. Each switching node 22 that an IP packet encounters on the path from host 20 to host 24 constitutes a separate hop. The IP packet has a header that contains a destination IP address. The destination IP address specifies host 24 as the destination. Each switching node 22 on the path between host 20 and host 24 uses the destination address in determining a next hop.

IP addresses were previously divided into three classes: Class A, Class B and Class C. The number of bits allocated to the network identifier 12 in the IP address and the number of bits allocated to the host identifier in the IP address was originally determined by the class of the IP address. With class A IP addresses, the host identifier was allocated three bytes; with class B IP addresses, the host identifier was allocated two bytes; and with class C IP addresses the host identifier was allocated a single byte. Many parties objected to this rigid bit allocation between host identifier and network identifier. As a result, a more flexible scheme was developed where masks were used to identify which bits in an IP address were allocated to the host identifier and which bits were allocated to the network identifier. A number of popular IP routing protocols utilize such masks.

Routers generally include routing tables to assist in forwarding IP packets to their proper destinations. The entries in the routing table hold forwarding information for IP address prefixes (i.e. portions of the IP addresses containing the most significant bits) for which routing information is known. For example, it may be known that all IP

packets destined to network 1.2 should be forwarded out over interface A of the node; thus, the entry encodes this knowledge.

Figure 4A shows an example of four routing table entries 30, 32, 34 and 36. Each routing table entry holds an address 40, a prefix length 42 and an interface 44. The address 40 field contains a prefix of an IP address. The prefix length 42 identifies the length of the prefix within the address field 40. For entry 30, the prefix is only a single byte (i.e. 8 bits) long. The interface 44 identifies the interface to which packets starting with the given prefix may be routed. The interface is a logical abstraction of a port (or other information) that identifies where a range of IP addresses (i.e. the addresses in the range defined by the prefix) should be directed.

Figure 4B shows an example of the topology of a portion of a computer network wherein the forwarding table entries 30, 32, 34 and 36 are utilized. In particular, node 50 has three interfaces: A, B and C. Interface C leads to network 1. The notation 1/8 in Figure 4B indicates that the IP address for the network has a prefix value of 1 and that it is 8 bits in length. Interface B leads to a portion of the computer network having IP addresses that start with the prefixes 1.2.3. Interface A leads to destinations having IP addresses that start with the prefix 1.2.4 and 1.2. Specifically, interface A leads to node 52, which, in turn, leads to the other destinations 1.2.4 and 1.2.

For each IP packet received by a node, the longest matching prefix found in the routing table is used to route the IP packet. Consider an IP packet that has a destination address of 1.2.4.7. In such an instance, entries 30, 34 and 36 contain matching prefixes for this IP address. Prefix 1.2.4, however, is the longest prefix and, thus, entry 34 is used to route the IP packet out interface A toward destination 1.2.4.

In conventional routers, the routing table is typically represented as a patricia tree. A patricia tree is a tree data structure that is used to simplify searching of the routing table. The patricia tree employs a binary representation of keys without storing keys in the nodes. Figure 5 shows an example of a portion of a patricia tree 60. Each node is associated with a particular portion of an IP address prefix. For example, the node b_0 is associated with bit 0 of an IP address prefix (i.e. the first bit in an IP address prefix). Each node may contain pointers to child nodes or to terminations. Furthermore each node may have a reference to a routing table entry for the prefix that the node represents. Each pointer leading from a node is associated with a bit value for the next bit in the prefix for the node. The structure is organized as a tree such that each level of the tree represents a successive bit sequence. Thus, node b_1 of Figure 5 represents a two bit sequence in the prefix where the first bit has a value of 0. The table entries are associated with the last node of the prefix. In the example of Figure 5, entry 62 for the

prefix 1, which is one byte in length (or 8 bits in length), is referenced by the node b_8 for the prefix bit sequence of "00000001." Similarly, entry 64 is referenced by the node b_{16}. Terminations, such as termination 65, are provided in the patricia tree to represent prefixes for which there is no associated forwarding table entry.

The patricia tree may also be implemented in a different fashion. The patricia tree may store the table entries so that the pointers point to the table entries (i.e. the table entries are in the tree as nodes). Hence, for a given node, a pointer associated.

The patricia tree provides a convenient search mechanism for conducting a binary search to identify whether any entries are associated with a particular prefix or portion of a prefix. One difficulty with the use of a patricia tree, however, concerns the number of memory accesses that must be performed to utilize the patricia tree. Addressing a node in the patricia tree requires a memory access. Thus, to search down to level 8 of the tree requires 8 memory accesses. Such memory accesses can be quite expensive in terms of time and computational overhead. Given that routers often handle extremely large volumes of IP packets, time and computational overhead are scarce resources that need to be conserved.

Summary of the Invention

The present invention addresses the limitation discussed above of conventional IP packet routing schemes by providing an packet forwarding approach that requires at most three forwarding table lookup accesses per destination address. By minimizing the number of lookup accesses, the present invention decreases computational overhead and the time required to determine how to properly route a packet. In one embodiment, the present invention uses three types of lookup arrays. A first type of lookup array is indexed by the first two bytes of a destination IP address for an IP packet. In some embodiments, the destination address is not used alone for the lookup; rather other fields such as the source address, destination port and source port are used in conjunction with the destination address during lookup. Nevertheless, it is worth considering the case wherein only the destination address is used.. The second type of lookup array is indexed by the third byte of the destination IP address. It contains entries for prefixes in the range of greater than two bytes and less than or equal to three bytes. Each entry in the first lookup array may have a separate associated second lookup array. If the second does not contain a matching entry, there are no entries that match the prefix formed by the first three bytes of the destination IP address; hence, the third type of lookup array must be used. The final byte of the destination IP address is used as an index to this table. A separate third lookup array may be provided for each entry in a second lookup

array. Thus, the lookup arrays are organized as a tree of lookup arrays in one embodiment of the present invention.

In accordance with one aspect of the present invention, a method is practiced in a digital logic device for forwarding data packets. The device includes a storage element having addressable storage locations. Multiple bits from header data for network layer packet are used as an index to locate a selected one of the storage locations. This selected storage location provides information regarding how the device should forward the IP packet. This information is utilized to forward the packet toward the destination.

In accordance with another aspect of the present invention, a first and a second forwarding lookup are provided in a device for forwarding an IP packet toward a destination, where the destination has a destination address comprising a sequence of bits. A prefix of multiple bits for the destination address is used as an index to locate a first entry in the first forwarding lookup. The first entry provides direction to the second forwarding lookup. The next sequential set of bits that follows the prefix of the destination address is used as an index to locate a second entry in the second forwarding lookup. The contents of the second entry are employed in forwarding the IP packet towards the destination address.

In accordance with a further aspect of the present invention, a forwarding lookup that has locations that are indexed by multiple bits is provided within a switch. The switch is in a network that employs a connectionless network protocol. For each data packet to be forwarded to a destination address, bits in the destination address are used to locate and access at least one location in the forwarding lookup. The location that is accessed is used to forward the data packet. Fewer locations are provided in the forwarding lookup than bits provided in the associated destination address.

In accordance with a further aspect of the present invention, a device for forwarding network layer packets to destinations (wherein the packets have associated header data) includes a first lookup structure. The first lookup structure holds entries that provide information regarding how to forward packets to their destinations. The entries are indexed by multiple bits. The device also includes a forwarding controller for using multiple bits from the header data as indices to locate entries in the first lookup structure. The forwarding controller also uses the entries in the first lookup structure in directing the forwarding of the packets to the destinations.

In accordance with another aspect of the present invention, a switch/router directs network IP packets towards destinations. The switch/router includes a first lookup array containing entries that are indexed by leading bits of destination addresses for IP packets. Each entry contains an instruction to assist in forwarding an IP packet

towards a destination. The switch/router also includes a second lookup array containing entries indexed by a successive set of bits that follow the leading bits in the destination addresses for IP packets. Each entry contains an instruction to assist in forwarding an IP packet towards a destination. The switch/router additionally includes a third lookup array containing entries indexed by a set of trailing bits that followed the successive set of bits in the destination addresses for IP packets. Each entry in the third lookup array contains an instruction to assist in forwarding an IP packet. The switch/router includes a forwarding engine for forwarding IP packets to the destinations. The forwarding engine accesses at least one entry in the lookup arrays that is indexed by a destination address for the IP packet being forwarded. The forwarding engine executes the instruction contained in the entry that is accessed.

Brief Description of the Drawings

An illustrative embodiment of the present invention is described below relative to the following drawings.

FIGURE 1 depicts the seven layers found in the OSI reference model.

FIGURE 2 depicts the major logical component of an IP address.

FIGURE 3 shows an example of a conventional computer network in which an IP address is employed.

FIGURE 4A depicts an example of conventional routing table entries.

FIGURE 4B depicts an example of a conventional computer network for which the routing table entries of Figure 4A are provided.

FIGURE 5 depicts an example of a portion of a patricia tree that is used to locate forwarding table entries in a conventional system.

FIGURE 6 is a block diagram illustrating the role of the switch/router in practicing the illustrative embodiment of the present invention.

FIGURE 7 is a block diagram illustrating major components of the switch/router of Figure 6.

FIGURE 8 depicts major components of a line card employed in the illustrative embodiment of Figures 6 and 7.

FIGURE 9 is a flow chart illustrating the steps that are performed in processing an incoming frame of data to properly forward an IP packet in the illustrative embodiment of Figures 6 and 7.

FIGURE 10 illustrates the manipulation of data in the illustrative embodiment of the present invention.

FIGURE 11 illustrates the format of a SONET frame.

FIGURE 12 illustrates the major components of an IP packet.

FIGURE 13 illustrates the format of header data used in IP forwarding lookup.

FIGURE 14 illustrates structures held in memory that are used in performing IP lookup in the illustrative embodiment of the present invention.

FIGURE 15 illustrates an interface structure.

FIGURE 16 illustrates a DANET structure.

FIGURE 17 is a flow chart illustrating the steps that are performed during IP lookup in accordance with the illustrative embodiment of the present invention.

FIGURE 18 depicts the use of a lookup element in the illustrative embodiment of the present invention.

FIGURE 19A illustrates an example where a first lookup array references a second lookup array.

FIGURE 19B illustrates an example where "smearing" is used so that a range of entries reference a common DANET structure.

FIGURE 19C illustrates an example where a first lookup array references an entry in the second lookup array, which references an entry in a third lookup array.

FIGURE 20 illustrates the logical format of a lookup element.

FIGURE 21 is a block diagram that illustrates the use of a rotor pointer and a TOS array pointer to obtain a destination handle for an IP packet.

Detailed Description of the Invention

The illustrative embodiment of the present invention provides a switch/router that forwards network layer packets toward their destination with fewer memory accesses on average during network layer forwarding lookup than conventional switching nodes. "Network layer packet" refers to a packet that complies with an OSI layer 3 protocol. Although the illustrative embodiment of the present invention will be described below for use with IP packets, the present invention may also be used for different types of network address lookup, such as with CLNP and other protocols. The switch/router employs a first forwarding lookup that may be indexed by the leading sixteen bits of the destination address for an IP packet. A second forwarding lookup is also provided within the switch/router. The second forwarding lookup may be indexed by the next successive eight bits in the destination address that follows the first sixteen bits. Lastly, a third forwarding lookup is provided in the switch/router. The third forwarding lookup may be indexed by the final 8 bits of the destination address of an IP packet. Entries in the third forwarding lookup are used when entries in the first

forwarding lookup and in the second forwarding lookup are not sufficient to forward the IP packet toward a destination.

The illustrative embodiment will be described relative to an implementation that uses IP, version 4. Nevertheless, those skilled in the art will appreciate that the present invention may also be practiced with other versions of IP, including version 6.

Analysis of IP packet addresses and traffic patterns reveals that the majority of IP packets only require a single lookup in the first forwarding lookup (i.e. most IP packets may be properly routed based on the first two bytes of their destination IP addresses). Thus, the majority of IP packets require only a single memory access. An overwhelming percentage of IP packets require only either a lookup in the first forwarding lookup or a lookup in both the first forwarding lookup and the second forwarding lookup. Thus, an overwhelming percentage of IP packets may be forwarded with only two memory accesses for IP lookup. As a result, the illustrative embodiment provides substantial time and computational savings.

In the illustrative embodiment, each lookup array entry or element contains an instruction. The instruction is executed by a lookup engine that is provided in the switch/router. The instruction tells the lookup engine what to do next during the lookup process. For example, an instruction in an element in the first forwarding lookup may instruct the lookup engine to access an element in the second forwarding lookup. The element that is accessed in the second forwarding lookup array may contain an instruction directing the lookup engine to use a particular data structure, that holds information regarding which output port to use in forwarding the IP packet.

The switch/router of the illustrative embodiment is presumed to be positioned in a computer network where IP packets need to be forwarded toward destinations. The switch/router of the illustrative embodiment is suitable for use in computer networks, such as, for example, the Internet, an intranet or an extranet. Figure 6 depicts the basic role of the switch/router 66 in the illustrative embodiment. In particular, an IP packet 64 enters the switch/router 66 via an input port 68. The switch/router 66 determines which output port 70 to use in outputting the IP packet 64 so as to ensure that the IP packet heads towards the desired destination. The IP packet 64 may be encapsulated into frames and may be enter the switch/router 66 along with other IP packets. The decision regarding how to forward the IP packet 64 within the switch/router 66 involves IP forwarding lookup, which will be described in more detail below.

Figure 7 illustrates an example of a portion of the basic layout for the switch/router 66. The components shown in Figure 7 may reside in a single box (i.e. housed within a single housing). The switch/router 66 is able to receive and process

multiple input data streams, concurrently. These input streams arrive at the switch/router 66 over separate links. In the illustrative embodiment these input data streams are SONET data streams(SONET is an acronym for synchronous optical networks). SONET is a standard that specifies a synchronous level one transport signal at 51.84 megabits per second. This standard defines a family of fiber-optic transmission rates that facilitates the internetworking of transmission products for multiple venders. The standard defines a physical interface, optical line rates known as Optical Carrier (OC) signals, and a frame format. The SONET optical line rates are defined as follows:

OC Level	Line Rates	Capacity
OC-1	51.84 Mbps	28 DS1s or 1 DS3
OC-3	155.52 Mbps	84 DS1s or 3 DS3s
OC-9	466.56 Mbps	252 DS1s or 9 DS3s
OC-12	622.08 Mbps	336 DS1s or 12 DS3s
OC-18	933.12 Mbps	504 DS1s or 18 DS3s
OC-24	1.244 Gbps	672 DS1s or 24 DS3s
OC-36	1.866 Gbps	1008 DS1s or 36 DS3s
OC-48	2.488 Gbps	1344 DS1s or 48 DS3s
OC-96	4.976 Gbps	2688 DS1s or 96 DS3s
OC-192	9.953 Gbps	5376 DS1s or 192 DS3s
OC-255	13.21 Gbps	

In the above table, DS refers to a known standardized hierarchy of digital signal speeds used to classify capacities of lines and trunks. The fundamental speed level is DS-0, which corresponds with 64 kilobits per second. DS-1 corresponds to 1.544 megabits per second, and DS 3 corresponds to 44.736 megabits per second.

Each line card 76, 78, 80 and 82 is designed to receive an OC-48 input stream, which corresponds to the 2.488 gigabits per second (Gbps). Multiplexers 72 and 74 are provided to multiplex four OC-12 input data streams in order to produce an OC 48 input data stream at line cards 82 and 76, respectively. In the example depicted in Figure 7, it is presumed that separate OC-48 input data streams are received by line cards 78 and 80, respectively.

The line cards 76, 78, 80 and 82 contain intelligence for receiving and transmitting IP packets. Each line card 76, 78, 80 and 82 is positioned on a common chassis within the switch/router 66. Each line card 76, 78, 80 and 82 contains at least one application specific integrated circuit (ASIC) 84, 86, 88 and 90 that performs the IP

forwarding lookup. Figure 8 depicts major components a line card 100 in more detail. The line card 100 includes a microprocessor 102 and memory 104. The line card 100 also includes an ASIC 106 that has a lookup engine 108. The lookup engine 108 may be implemented in a number of different forms, including as a separate processor. Although the ASIC provides a hardware implementation for IP forwarding lookup, those skilled in the art will appreciate that the present invention also encompasses a software implementation. Other ASICs may be provided on the line card 100 to implement other functionality.

The ASIC 84, 86, 88 and 90 on each line card 76, 78, 80 and 82 is responsible for receiving incoming IP packets, determining the appropriate destination handle for the IP packets and passing the IP packets over the interconnect to the appropriate output line card. The destination handle specifies to the output line card how the IP packet should be forwarded. The interconnect 92 is a interconnection fabric that interconnects the line cards 76, 78, 80 and 82. A control processor 94 oversees and manages operations within the portion of the switch/router 66 shown in Figure 7.

Those skilled in the art will appreciate that the present invention need not be practiced with a switch/router configuration like that shown in Figures 6 and 7. The depiction shown in Figure 6 is intended to be illustrative and not limiting of the present invention. For example, the IP forwarding could be performed in a computer system, such as a personal computer. Moreover, the IP forwarding lookup need not be performed by an ASIC but rather may be performed by a dedicated forwarding microprocessor or by a state machine. As mentioned above, the IP forwarding lookup may be implemented solely by software. In addition, the intelligence need not reside at the line cards but rather claims line cards may be used with an intelligent processor performing the IP forwarding lookup. Still further, the switch/router need not have four line cards but rather may have a different number of line cards. The input data need not be SONET streams holding data in SONET frames. Other types of data formats and streams may be received in practicing the present invention.

An example is helpful to illustrate operation of the switch/router 66 in forwarding an IP packet. Suppose that an IP packet is received by SONET multiplexer 74. The IP packet is then received by the line card 76 and processed by the ASIC 84. The ASIC 84 directs the IP packet over the interconnect 92 to line card 82. Line card 82 subsequently directs the IP packet out towards SONET mux 72 so that the IP packet may be output toward the appropriate destination.

Figure 9 is a flow chart that provides an overview of the processing performed on data that is received by the switch/router 66. It is presumed that this data contains at

least one IP packet. Initially, the data start off in state 128 (Figure 10) where a SONET frame 130 is received from one of the links. The SONET frame 130 encapsulates a frame of data that is transmitted in the format identified by the SONET standard. Figure 11 provides a block diagram illustrating the format of a SONET frame 130. A SONET frame 130 includes 90 octets (8 bit bytes) across and 9 rows down. The payload is contained in the synchronous payload envelope (SPE). The SPE contains 9 bytes that are dedicated to path overhead (OH). The SONET frame 130 also contains section overhead 146 and line overhead 148. The section overhead 146 and line overhead 148 are part of the SONET transport overhead. In this context, "overhead" refers to header information that is provided for various layers of the computer network.

As can be seen in Figure 10 the SONET frame 130 encapsulates a layer two structure (i.e. a structure provided by a layer two protocol, where layer two is defined by the OSI model). At least one IP packet 134 is held within the SONET frame 130 and the level two structure 132. The SONET frame is then decapsulated by the switch/router 66 (step 112 in Figure 9). The switch/router 66 contains hardware that is designed for decapsulating the SONET frame. After decapsulation, the layer two structure 132 that contains the IP packet 134 is exposed (as indicated by state 136 in Figure 10).

The switch/router 66 then peels open the layer two structure 134 by removing the layer two header so as to gain access to one or more IP packets 134 (step 114 in Figure 9). The layer two structure may be, for example, a point-to-point protocol (PPP) frame, an ATM cell or a frame relay frame.

The lookup engine 108 of the ASIC 106 obtains a single IP packet from the layer two structure (step 116 in Figure 9). The ASIC 106 knows that the layer two structure contains IP packets based upon interface information. The switch/router 66 maintains interface information regarding interfaces in which incoming data is received. Each interface is associated with a particular line card and port. The interface information identifies the nature of the data that is being received. For instance, the data may be identified as containing IP packets.

The IP header 152 (Figure 10) from the IP packet 134 is copied along with some port information 141 from the transport header 143 to produce header data 153 (step 118 in Figure 9). As shown in Figure 12, the IP packet 134 includes a header 152 and data 154. Thus, in step 119, the data being processed transitions from state 138 to state 139 (see Figure 10).

Figure 13 shows an example of the header data 152 that is used for IP forwarding lookup for IP version 4. All of the fields in the header data 153 other than fields 184 and 186 (which are copied from the transport header 143) are copied from the IP header 152.

The header data 153 includes a version field 160 that holds information regarding the version of the IP protocol being used. For version 4 IP packets, this field 160 holds a value of 4. The Internet header length (IHL) field 162 identifies the length of the header in multiples of 4 octets. The differential services (DF) 164 holds a number that identifies a particular handling or treatment for the packet. The total length field 166 holds information regarding the total length of a packet before any fragmentation occurs. The identification field provides an identification value for the packet that may be used if the packet is later fragmented to associate the fragments with the original packet.

The header data 153 includes flags 170. The flags 170 include a DF flag and a MF flag. the DF ("don't fragment") flag indicates whether a datagram (carried at least in part by the packet) is to be fragmented. The MF ("more fragment") flag identifies whether there are more fragments or whether the packet holds the last fragment of the datagram. The fragment offset field 172 holds an offset value that identifies the offset at which the fragment belongs to the reassembled packet. The time to live field 174 identifies the time period for which the packet is valid and after which the packet should be discarded. The protocol field 176 holds a value that allows the network layer of the destination end node to know which protocol running within the end node should receive the packet. A header check sum field 178 is provided. The header data 153 also includes a source address 180 that identifies the source for which the packet originated. A destination address field 182 holds a destination address for the destination to which the IP packet is to be forwarded. The header data 153 also includes a source port field 184 and a destination port field 186 that are copied from the transport header for identifying respective ports. The port fields 184 and 186 may be used in quality of service (QOS) processing or in other fashions, such as access control filtering.

Once the header data 153 has been gathered, the destination address 182 may be used to perform a forwarding lookup within the forwarding lookup arrays (step 120 in Figure 9). The IP forwarding lookup need not rely solely on the destination address; rather additional fields in the header data 153 may be used in conjunction with the destination address. For instance, the source address, source port and destination port may be used along with the destination address. These other fields may be used in providing certain QOS. For purposes of simplicity, the discussion below will initially focus on the instance wherein only the destination address is relied upon. This lookup identifies where the IP packet should be output by the switch/router 66. The lookup process will be described in more detail below. The IP packet then is forwarded across the interconnect 92 to a line card so as to be output from the switch/router 66 (step 122 in Figure 9).

Packets may require a QOS processing or not. Packets that require a QOS processing may be subject to a different QOS than other types of packets. QOS processing evaluates additional fields in the header data 153 to determine whether a packet is to be classified into a specific QOS flow or not. A QOS processing identifies these flows and segregates such packets for special processing. A filter specification identifies what fields are evaluated and the values the fields should have for a given type of QOS. As a result, certain packets may be routed based upon the fields evaluated during QOS processing rather than based upon the destination address alone. The destination address case is discussed here as the basic approach and may be used in conjunction with QOS processing to determine how to forward an IP packet.

In performing the forwarding lookup, the lookup engine 108 uses a number of internal structures, including tables, arrays and other data structures. Figure 14 depicts several of the major varieties of structures that are utilized during a forwarding lookup for IP packets. Interface structures 210 contain information regarding interfaces. An interface generally refers to a link with another switching node in a computer network. Figure 15 shows an example of an interface structure 210 for a given interface. The interface structure 210 also contains an initial lookup element 220. The initial lookup element 220 is an array lookup element that contains an initial instruction that is executed at the beginning of forwarding lookup for an IP packet. The use of this initial lookup element 220 will be described in more detail below. The interface structure 210 may also contain a number of counters 221 that hold counts which are useful in gathering statistics regarding traffic over the interface.

The forwarding lookup also uses lookup arrays 212 composed of lookup elements. The format and use of these lookup elements will be described in more detail below. The forwarding lookup may also access a SANET 214 or a DANET 216. A SANET 214 is a data structure that holds a number of structures for respective source addresses. The structures hold useful information regarding source addresses that may be exploited for QOS and TOS. The DANET 216 holds DANET structures that contain information regarding destination addresses that is used in next hop determination. The DANET structures have a format like that shown in Figure 16. In particular, each DANET structure 222 holds a field 224 that may contain a destination handle, a pointer to a TOS array or a pointer to a rotor. As mentioned above, a destination handle is a composite data structure that holds useful information regarding where a given IP packet should be directed so that it is properly output towards a destination. The switch/router 66 uses the destination handle on the transmission side to determine where to send an IP packet (i.e. what line card and output port should be used). Field 224 may instead

contain a pointer to a rotor that contains a set of destination handles or a pointer to a type of service (TOS) array that holds a set of destination handles. The destination handles in the TOS array are indexed by a TOS parameter. The DANET structure 222 contains a number of counters 225 including packet counters and byte counters. These counters 225 are useful in monitoring traffic to a destination and may be used in QOS processing. The DANET structure 222 may also contain other data 226.

Figure 17 provides a flow chart of the steps that are performed during best-effort forwarding lookup for a unicast IP packet. The lookup determines how to send the IP packet to the next hop toward the destination. The switch/router 66 knows the interface on which the IP packet arrived. The interface structure for the associated interface is accessed, and the lookup engine 108 processes the initial lookup element contained in the interface structure (step 230 in Figure 17). As shown in Figure 15, the interface structure 210 includes a lookup array element 220 that contains an instruction. The instruction in an array lookup instruction which identifies the array to which the lookup is to be applied. The lookup element 220 (Figure 18) contains an opcode 256 for array lookup. The lookup element 220 also contains an array address 252 and a header nibble select 254. A nibble is 4 bits, and different nibbles within the header may be utilized to generate an index to an array lookup element in a lookup array. Information in the header, other than the destination address may be used for lookup, and the header nibble select 254 identifies what information to use for lookup. The array address 252 identifies the location of the lookup array 264 and may be combined with the 16 address bits 260 to locate the lookup element 266 within the lookup array 264. Thus, initially the entry 266 in the first lookup array 264 is accessed and processed (step 232 in Figure 17).

As shown in Figure 20, this lookup element in the first forwarding lookup array 264 contains an array address, header nibble select and opcode. The opcode may direct the lookup engine 108 to another forwarding lookup array. Hence, the next successive lookup array must be accessed. Figure 19A shows an example wherein a lookup element 272 in lookup array 264 identifies an array address for a second lookup array 274. The second lookup array 274 is indexed by the third byte within the destination address. The lookup elements in the second lookup array 274 include lookup elements 276 and 278 for the prefixes 1.2.3 and 1.2.4, respectively.

Figure 19A also shows an example wherein a lookup element 273 contains an opcode that directs the lookup to a different eight bit lookup array 275. The third byte of the destination address is used as an index into this eight bit lookup array 275 to locate a lookup element 277. As was mentioned above, the lookup arrays are organized as a tree

with the top level of the tree containing references to the next level of the tree. Hence, there may be a significant number of eight bit lookup arrays referenced by the sixteen bit lookup array in the implementation described for the illustrative embodiment of the present invention.

The above discussion has assumed that the instruction in the lookup element contained in the first lookup array is an array lookup instruction for a second lookup array. In some instances, the first lookup element may contain a set DANET instruction that associates a given DANET structure with the IP packet. This DANET structure contains a destination handle, or a pointer to a rotor or a TOS array from which a destination handle may be derived. In such a case, the lookup element is associated with a prefix that matches the first 16 bits of the destination address for the IP packet and the known forwarding information may be employed to forward the IP packet. Such a set DANET instruction may be found at any of the different layers of tables of forwarding lookup arrays, depending on where a match is found.

Multiple lookup elements may reference the same next level array or may references the same DANET structure. Figure 19B shows an instance wherein an eight bit lookup array 278 is referenced by a lookup array element 279 in the sixteen bit lookup array 264. The eight bit lookup array 278 contains 256 entries corresponding to the 256 possibilities of possible values that the third byte of the destination address may assume. The entries in the range for prefixes 1.2.128 through 1.2.255 all point to DANET structure 282 as the DANET structure to be used, except for the entry for 1.2.129. The entry for 1.2.129 indicates that a different DANET structure 280 is to be utilized. DANET structure 280 is for prefix 1.2.129/25 and DANET structure 28 is for prefix 1.2.1.128/17. This smearing provides an optimization so that a large number of copies of a given DANET structure need not be utilized, and, thus, the smearing saves storage space. This approach also accounts for instances wherein the matching prefix is between 17 and 23 bits in length.

In step 234, the lookup engine 238 determines whether it is done or not. The instruction that is executed by the lookup engine in step 232 will inform the lookup engine whether it is done or not. Where a match is found, the DANET structure that is set by the set DANET instruction is used in forwarding the packet and IP lookup is complete (step 242 in Figure 17). In other instances, there is no matching prefix of 16 bits or less and the lookup must continue with the second forwarding lookup array, which contains 2^8 elements and is induced by the third byte of the destination address.

If the lookup engine 108 is directed to look to the second forwarding lookup array, the lookup engine accesses the appropriate lookup element and the second lookup

array then processes the entry (step 236 in Figure 17). This lookup element may contain an instruction of the same variety of those discussed above relative to the first lookup array. In step 238, the lookup engine 108 determines whether it is done or not. If the lookup engine is not done, the instruction that was processed advises the lookup engine to look to the third lookup array to determine how to process the IP packet. This means that there was no matching prefix of 24 bits in length or less. Hence, the third and final forwarding lookup array containing 2^8 entries is to be accessed. As such, the lookup engine 108 accesses a lookup element in the third lookup array and processes the element (step 240 in Figure 17). The identified DANET structure is then used in forwarding the packet (step 242 in Figure 17).

Figure 19C shows an example where lookup elements from all three levels of the forwarding lookup arrays are utilized. In particular, a lookup element 272 in the 16 bit or first lookup array 264 is processed and directs the lookup engine 108 to access lookup element 282 in the second lookup array 274. The lookup element 282 contains an instruction to perform an array lookup on lookup element 292 within the third forwarding lookup array. The instruction in lookup element 292 is executed to set the appropriate DANET structure so that it is associated with the IP packet.

As was mentioned above, the DANET structure that is employed for use in forwarding an IP packet need not directly include the destination handle 215 (see Figure 21) but rather may include an indirect reference to obtaining the destination handle. For example, The DANET structure 222 may include a field that contains a pointer to a TOS array 310, which is a destination handle array. The TOS array 310 is indexed by a TOS parameter. The TOS offered to a packet may vary and may be expressed as a TOS parameter value. This value may be taken from field 164 of the header date 153, for example. The TOS parameter value acts as an index to the TOS array 310 to select a destination handle for the IP packet. The DANET structure 222 may also contain a reference to a rotor 314 that, in turn, references a destination handle 315. The TOS array 310 may also contain a reference to a rotor 314 rather than a direct reference to a destination handle 314. The rotor 314 is a structure that contains a set of destination handles and is used in the illustrative embodiment to facilitate aggregation of multiple lower speed links to present a virtual higher speed link. The rotor leg (i.e. which entry in the rotor is used) may be programmatically selected by either a randomly generated index or based on a hash of the fields that identify the QOS flow for the packet.

While the present invention has been described with reference to an illustrative embodiment thereof, those skilled in the art will appreciate that various changes in form and in detail may be made without departing from the intended scope of the present

invention as defined in the appended claims. For example, a different number of lookup arrays may be used and the lookups need not be arrays but may be organized differently such as lists, tables, etc. Furthermore, the arrays need not be indexed along byte boundaries. For instance, the first forwarding lookup array may be indexed by 15 bits rather than 16 bits. In addition, the array elements need not include instructions but rather may contain data or pointers.

Claims

1. In a device for forwarding data packets, having a storage containing storage locations, a method comprising the steps of:

providing header data for a destination of a network layer packet;

using multiple bits from the header data as an index to locate a selected one of the storage locations that provides information regarding how the device should forward the network layer packet; and

employing the information provided by the selected storage location to forward the network layer packet toward the destination.

2. The method of claim 1 wherein the selected storage location contains an instruction regarding how the device should forward the network layer packet and wherein the instruction is executed in the employing step.

3. The method of claim 1 wherein more than a byte from the destination address is used as the index.

4. The method of claim 1 wherein the network layer packet contains a header and wherein the method further comprises the step of extracting information address from the header.

5. The method of claim 1 wherein the packet is an IP packet.

6. In a device for forwarding an Internet Protocol (IP) packet toward a destination having a destination address containing a sequence of bits, a method comprising the steps of:

providing a first forwarding lookup and a second forwarding lookup;

using a prefix of multiple bits from the destination address of the IP packet as an index to locate a first entry in the first forwarding lookup;

where the first entry in the first forwarding lookup provides direction to the second forwarding lookup, using a next sequential set of bits following the prefix in the destination address as an index to locate a second entry in the second forwarding lookup, said second entry having contents; and

employing the contents of the second entry in forwarding the IP packet toward the destination address.

7. The method of claim 6 wherein the step of employing the contents of the second entry comprises executing an instruction contained in the second entry to forward the IP packet toward the destination address.

8. The method of claim 6 wherein the first entry contains an instruction to use the second forwarding lookup.

9. The method of claim 6 wherein the method further comprises the step of providing a third forwarding lookup and wherein the step of employing the contents of the second entry comprises identifying that the third forwarding lookup should be used in forwarding the IP packet.

10. The method of claim 9 wherein the method further comprises the steps of employing a final sequential set of bits in the destination address following the next sequential set of bits in the destination address as an index to locate a third entry in the third forwarding lookup and employing the contents of the third entry in forwarding the IP packet toward the destination address.

11. The method of claim 6 wherein the device includes an application specific integrated circuit (ASIC) and wherein the ASIC performs the steps of the method.

12. In a switch having a memory in a network that employs a connectionless network protocol, a method of forwarding a data packets, each having an associated destination address comprising the steps of:

providing a forwarding lookup with locations in the memory, wherein the locations are indexed by multiple bits; and

for each data packet to be forwarded, employing bits in the destination address to locate and access at least one location in the forwarding lookup to forward the data packet, wherein fewer locations in the forwarding lookup are located and accessed than bits in the associated destination address.

13. The method of claim 12 wherein the data packets are Internet protocol (IP) packets.

14. The method of claim 12 wherein at most three locations in the forwarding lookup are used to forward any of the data packets.

15. A device for forwarding network layer packets to destinations wherein the packets include header data, comprising:

a first lookup structure holding entries that provide information regarding how to forward network layer packets to their destinations, said entries being indexed by multiple bits; and

a forwarding controller for using multiple bits from the header data as indices to locate entries in the first lookup structure and for using the entries in the first lookup structure in directing the forwarding of the network layer packets to the destinations.

16. The device of claim 15 wherein the forwarding controller includes a processor for executing instructions and wherein the entries on the first forwarding lookup structure includes instructions to be executed by the processor to provide information regarding how to forward network layer packets to their destinations.

17. The device of claim 15 wherein the entries in the first lookup structure are indexed by more than a byte of bits.

18. The device of claim 17 wherein the entries in the first lookup structure are indexed by two bytes.

19. An apparatus for forwarding IP packets, comprising:

multiple lookup structures for assisting in the forwarding of the IP packets; and

a processor for executing the instructions in the entries of the lookup structures to forward the IP packets.

20. The apparatus of claim 19 wherein the apparatus is an application specific integrated circuit (ASIC).

21. A switch/router for directing IP packets toward destinations, comprising:

a first lookup array containing entries indexed by leading bits of destination addresses for IP packets, each entry containing an instruction to assist in forwarding an IP packet towards a destination;

a second lookup array containing entries indexed by a successive set of bits that follow the leading bits in the destination addresses for IP packets, each entry containing an instruction to assist in forwarding an IP packet towards a destination;

a third lookup array containing entries indexed by a set of trailing bits that follow the successive set of bits in the destination addresses for IP packets, each entry containing an instruction to assist in forwarding an IP packet; and

a forwarding engine for forwarding IP packets to destinations, wherein for each IP packet being forwarded, said forwarding engine accesses at least one entry in the lookup arrays indexed by a destination address for the IP packet being forwarded and executing the instruction contained in the entry that is accessed.

22. The switch/router of claim 21 further comprising input ports and interface structures that hold information regarding the input ports on which IP packets arrive.

23. The switch/router of claim 22 wherein the interface structures contain instructions for directing the forwarding engine to access the first lookup array.

24. In a device for forwarding data packets wherein the device includes a storage having storage locations, a computer-readable medium holding computer-executable instructions for performing a method, comprising the steps of:

using multiple bits from header data for an network layer packet as an index to locate a selected one of the storage locations that provides information regarding how the device should forward the network layer packet; and

employing the information provided by the selected storage location to forward the network layer packet toward the destination.

25. The computer-readable medium of claim 24 wherein the selected storage location contains an instruction regarding how the device should forward the network layer packet and wherein the instruction is executed in the employing step.

26. The computer-readable medium of claim 24 wherein more than a byte from the destination address is used as the index.

27. The computer-readable medium of claim 24 wherein the network layer packet contains a header and wherein the method further comprises the step of extracting the information from the header.

28. The computer readable medium of claim 24 wherein the packet is an IP packet.

29. In a device for forwarding an Internet Protocol (IP) packet toward a destination having a destination address composed of a sequence of bits, said device including a first forwarding lookup and a second forwarding lookup, a computer-readable medium holding computer-executable instructions for performing a method, comprising the steps of:

using a prefix of multiple bits from the destination address of the IP packet as an index to locate a first entry in the first forwarding lookup;

where the first entry in the first forwarding lookup provides direction to the second forwarding lookup, using a next sequential set of bits following the prefix in the destination address as an index to locate a second entry in the second forwarding lookup, said second entry having contents, and

employing the contents of the second entry in forwarding the IP packet toward the destination address.

30. The computer readable medium of claim 29 wherein the step of employing the contents of the second entry comprises executing an instruction contained in the second entry to forward the IP packet toward the destination address.

31. The computer-readable medium of claim 29 wherein the first entry contains an instructions to use the second forwarding lookup.

ABSTRACT

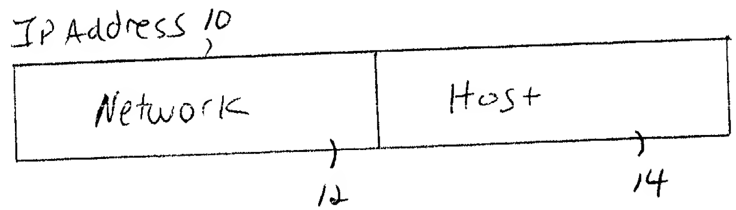
A switch/router contains intelligence for more quickly determining a next hop for an network layer packet. A network forwarding lookup table or array structure is configured so as to minimize the number of memory accesses required. This results in a decrease in time due to memory access and a decrease in computational overhead due to the memory accesses. In one embodiment, a first forwarding lookup is indexed by the first 16 bits of a destination address. A second forwarding lookup is indexed by the subsequent 8 bits of the destination address, and a final third forwarding lookup is indexed by the last 8 bits of the destination address. Each entry within a forwarding lookup contains either direction as to how properly forward the packet or reference to a next subsequent forwarding lookup.

OSI Layers

7	Application
6	Presentation
5	Session
4	Transport
3	Network
2	Data Link
1	Physical

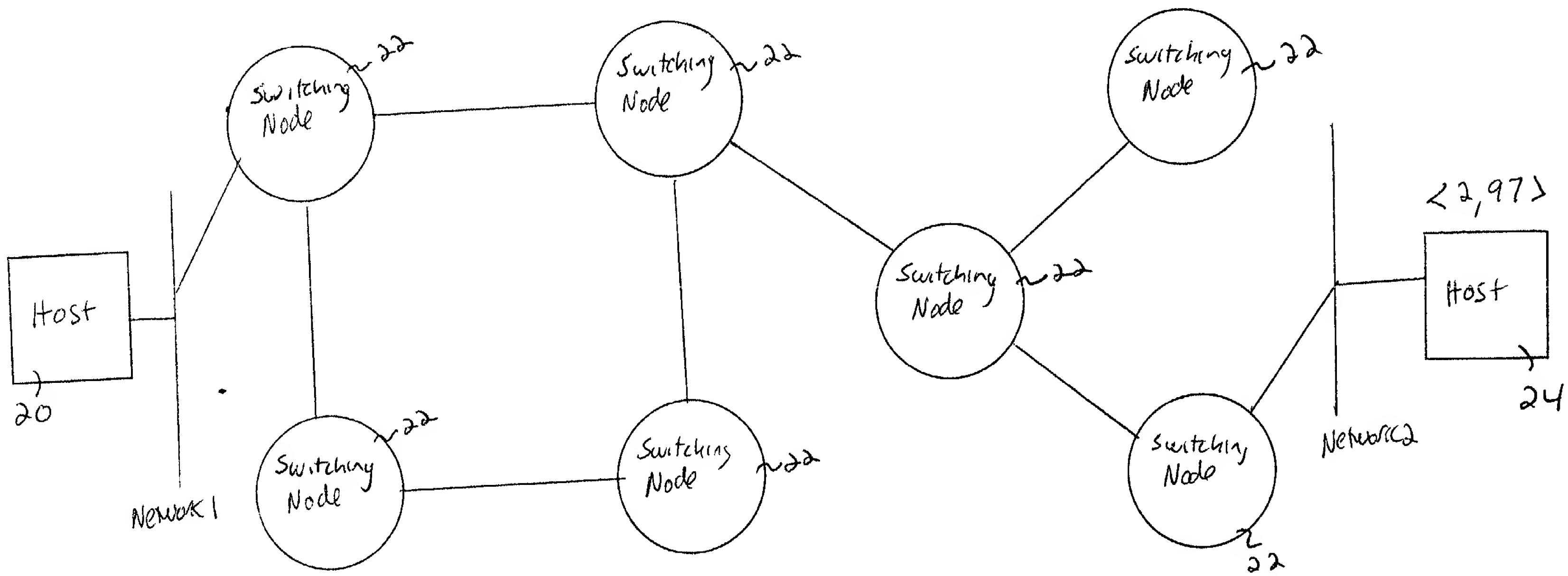
PRIOR ART

Figure 1



PRIOR ART

Figure 2



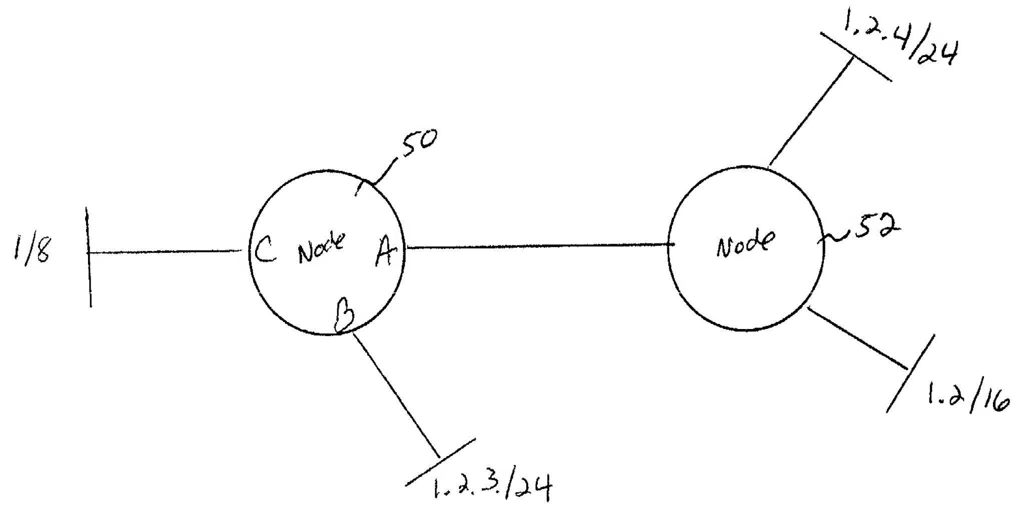
PRIOR ART

Figure 3

	Address 40	Prefix Length 42	Interface 44
30	1.0.0.0	8	C
32	1.2.3.0	24	B
34	1.2.4.0	24	A
36	1.2.0.0	16	A

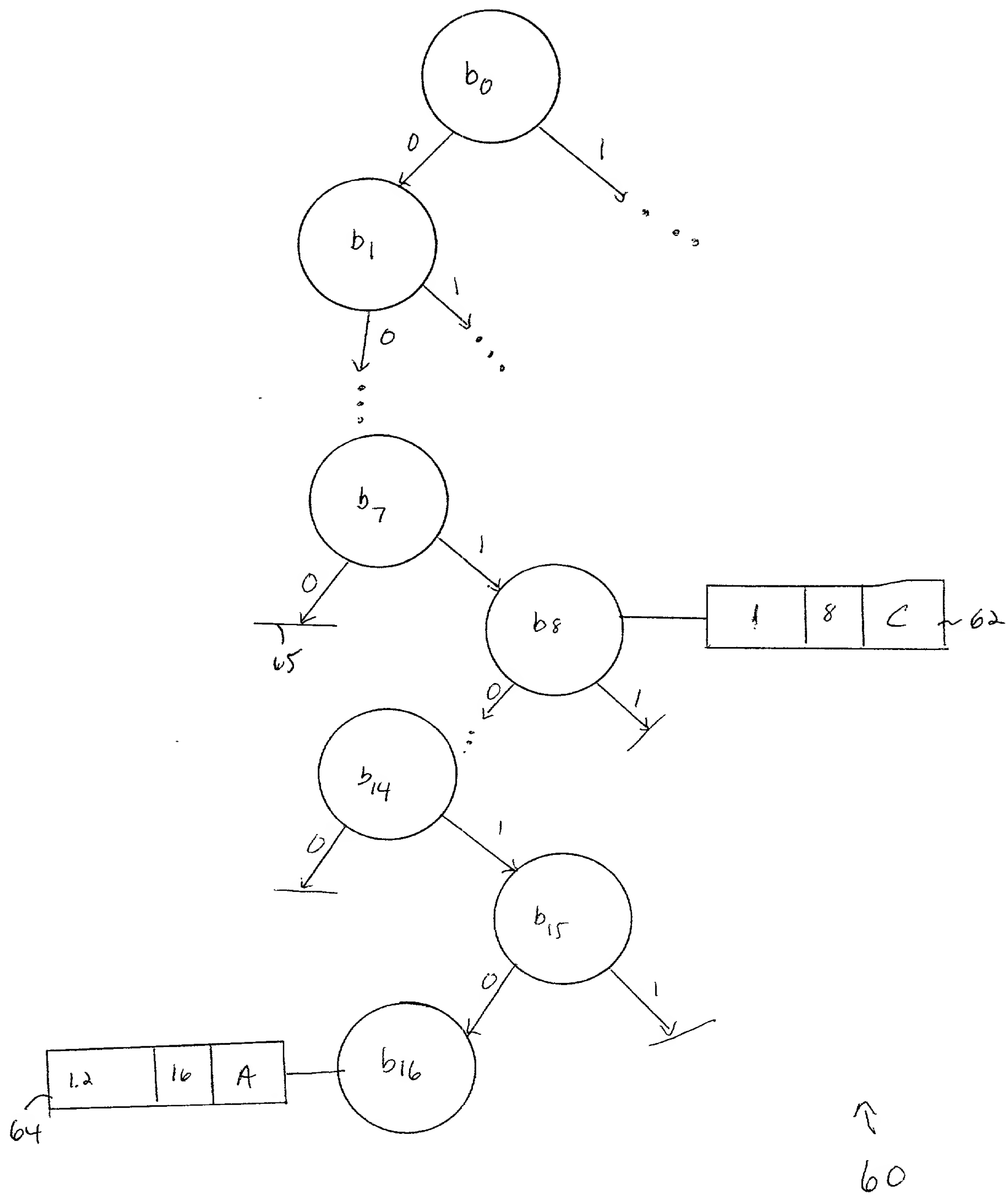
PRIOR ART

Figure 4A



PRIOR ART

Figure 4B



Prior Art
Figure 5

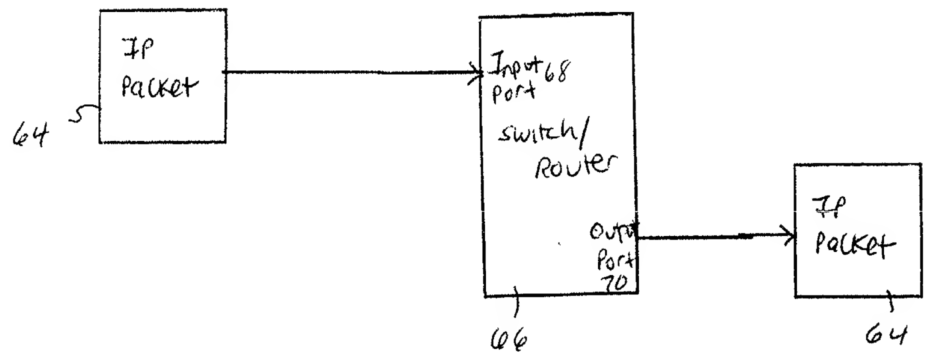

IP Packet
64
Input Port 68
switch/ Router
Output Port 70
66
IP Packet
64

Figure 6

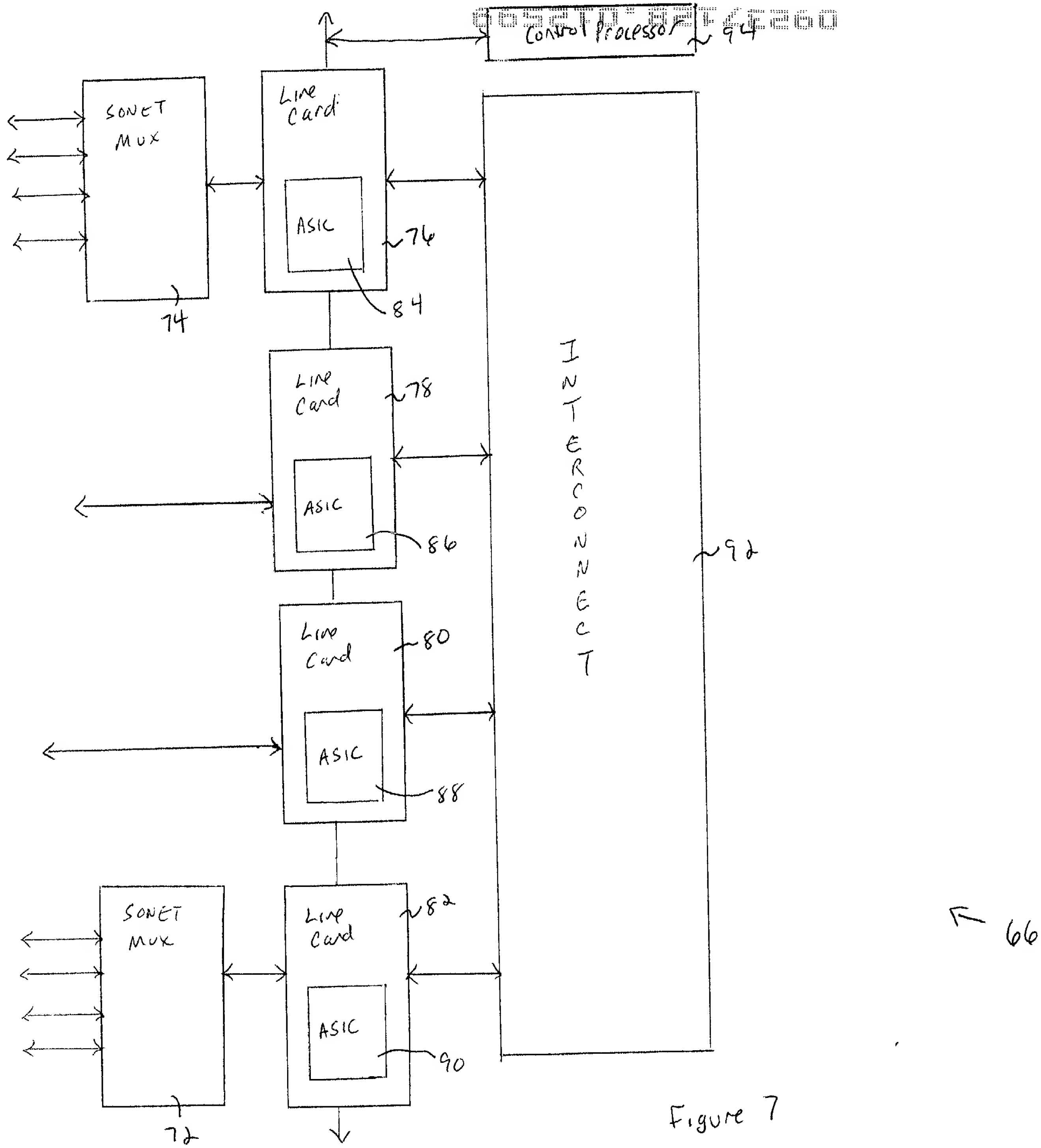

Control Processor
94
SONET MUX
74
Line Card
ASIC
76
84
Line Card
ASIC
78
86
Line Card
ASIC
80
88
SONET MUX
72
Line Card
ASIC
82
90
INTERCONNECT
92
66

Figure 7

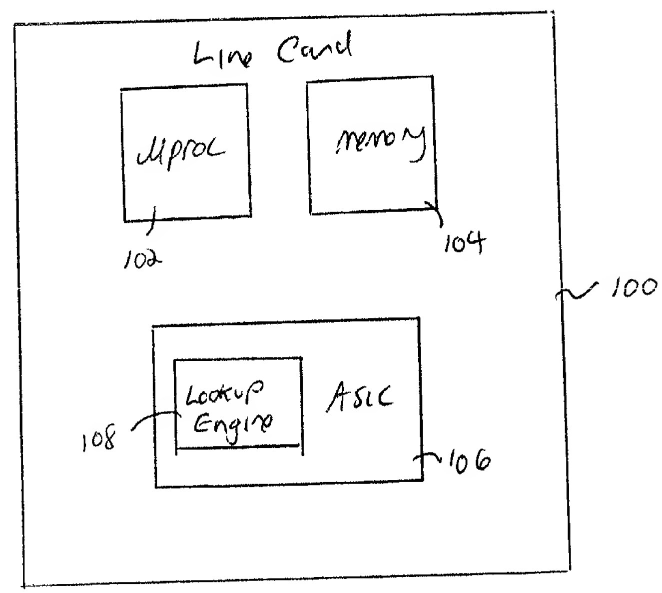

Line Card
μproc
memory
102
104
100
Lookup Engine
ASIC
108
106

Figure 8

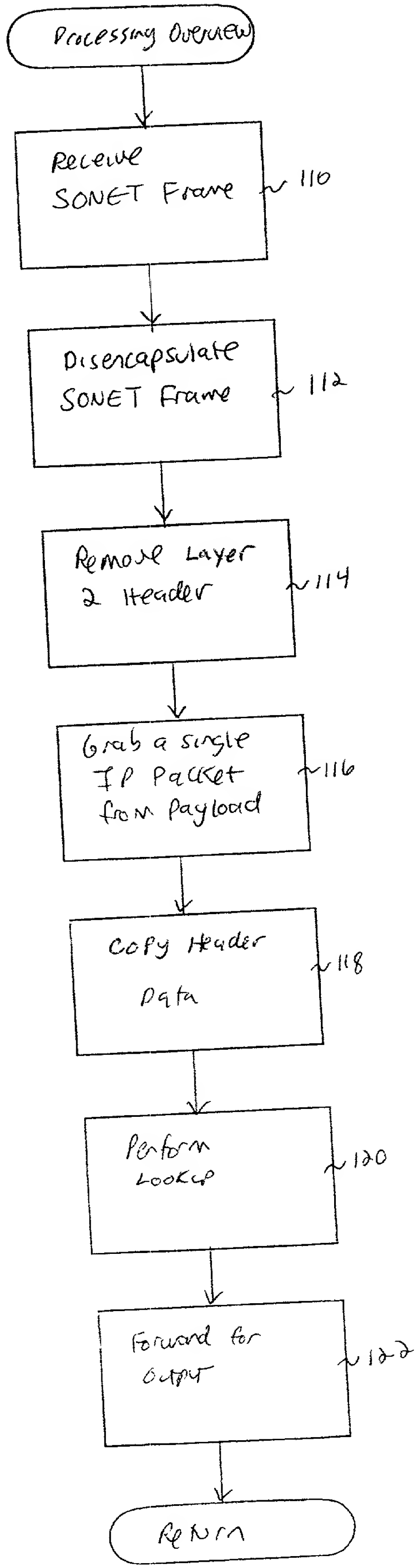

Processing Overview
Receive SONET Frame
110
Disencapsulate SONET Frame
112
Remove Layer 2 Header
114
Grab a single IP Packet from Payload
116
Copy Header Data
118
Perform Lookup
120
Forward for Output
122
Return

Figure 9

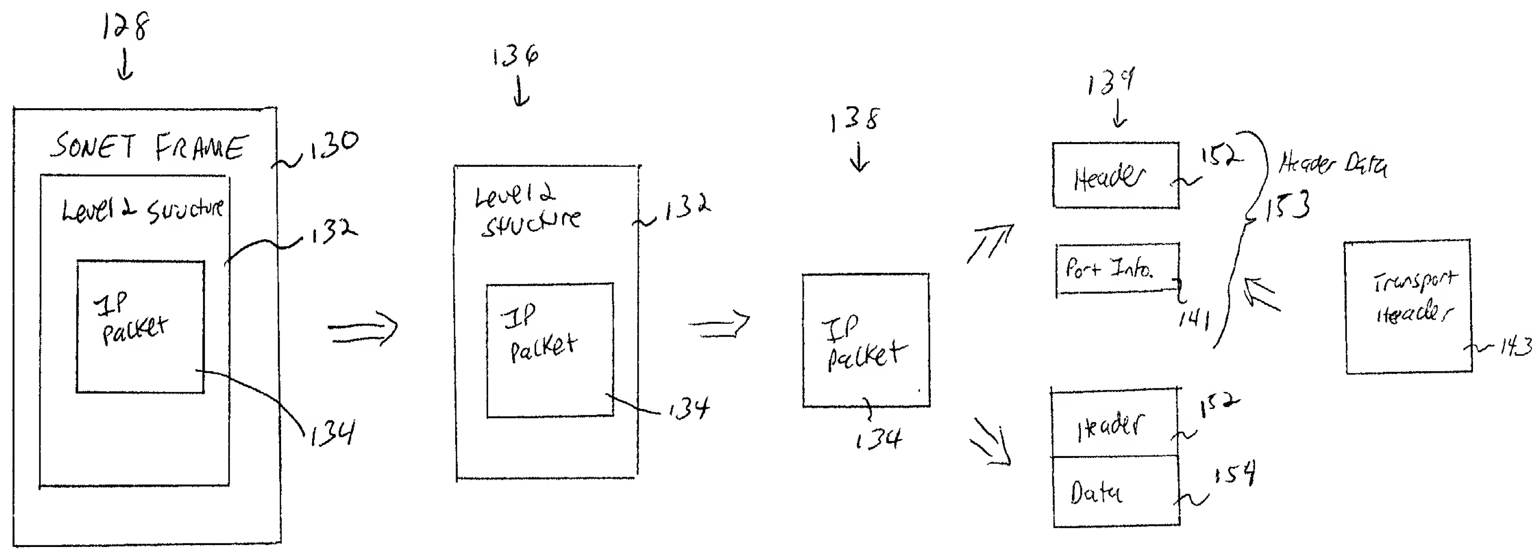

128
SONET FRAME
130
Level 2 Structure
132
IP Packet
134
136
Level 2 Structure
132
IP Packet
134
138
IP Packet
134
139
Header
152
Header Data
153
Port Info.
141
Transport Header
143
Header
152
Data
154

Figure 10

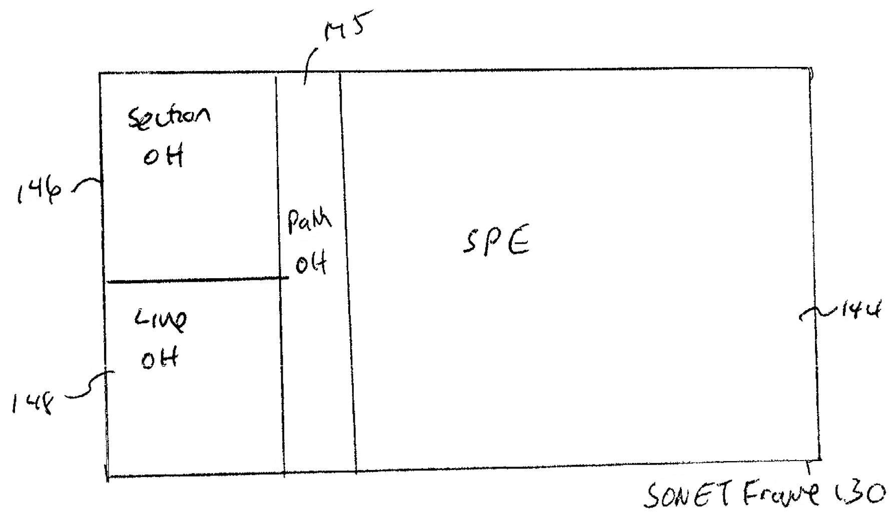


Figure 11

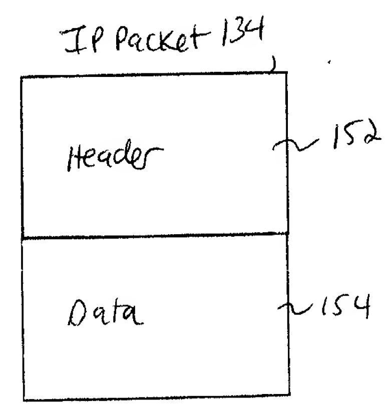
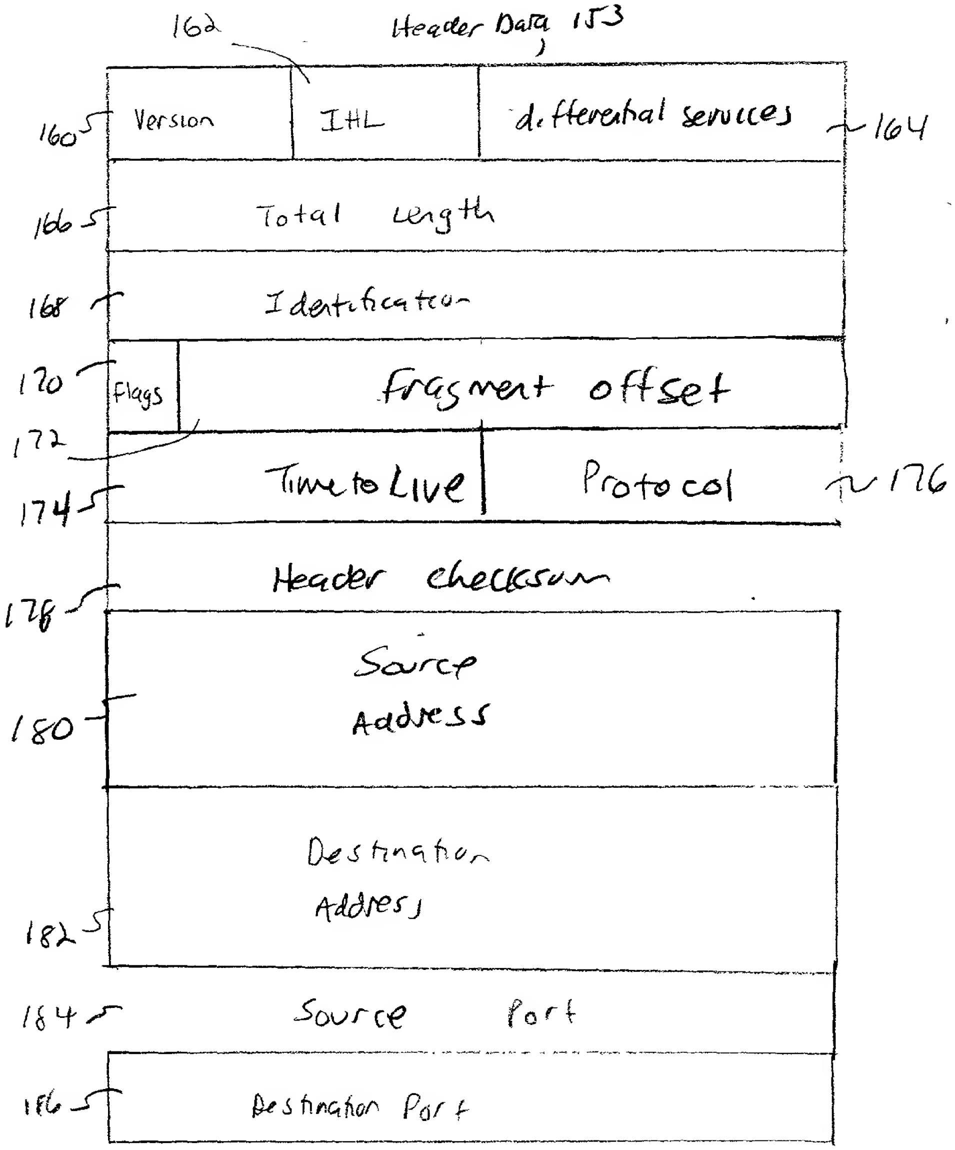


Figure 12



162
Header Data 153
Version
IHL
differential services
160
164
166
Total length
168
Identification
170
Flags
Fragment offset
172
174
Time to Live
Protocol
176
Header checksum
178
Source Address
180
Destination Address
182
184
Source Port
186
Destination Port

Figure 13

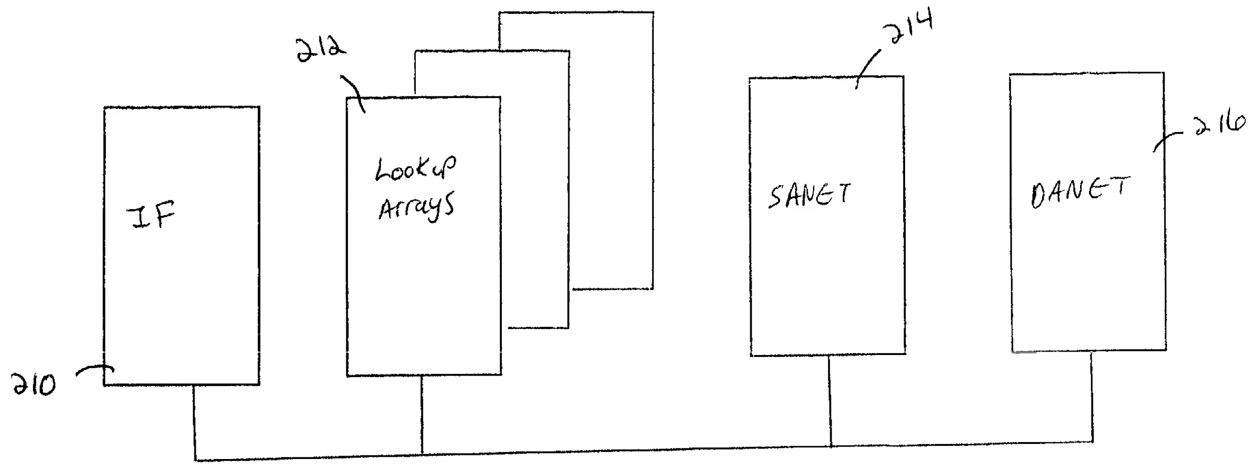
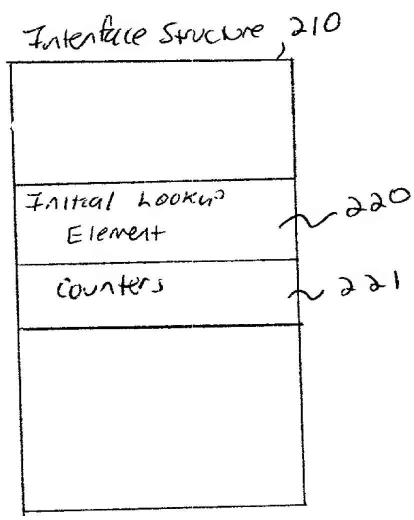


Figure 14



Interface Structure
210
Initial Lookup Element
220
Counters
221

Figure 15

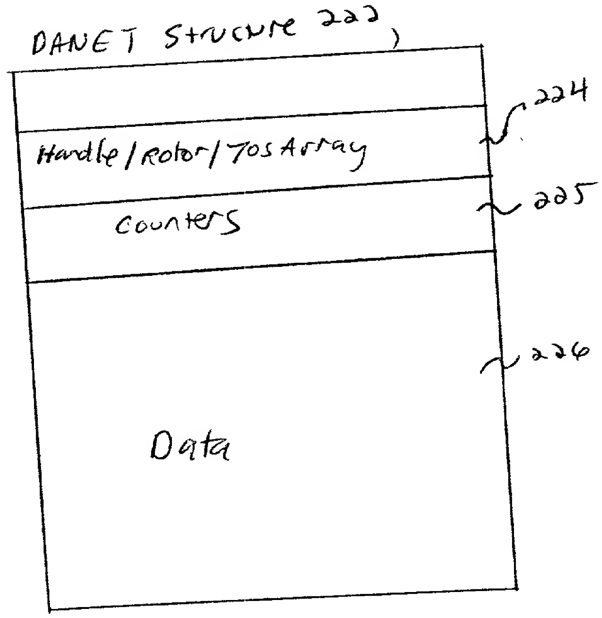
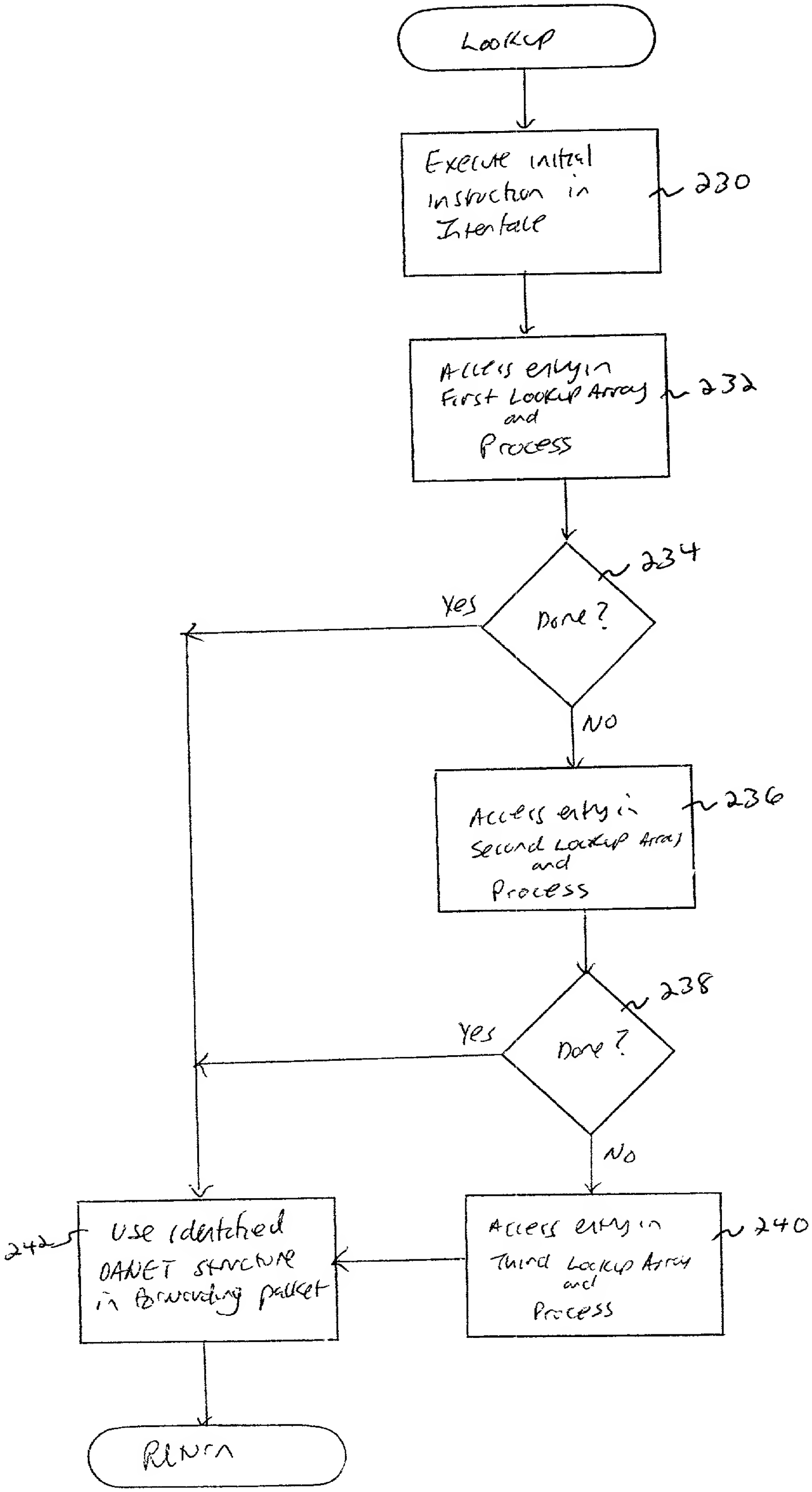

DANET Structure 222
224
Handle/Rotor/TOS Array
225
Counters
226
Data

Figure 16



Lookup
Execute initial instruction in Interface
230
Access entry in First Lookup Array and Process
232
Done?
234
Yes
NO
Access entry in Second Lookup Array and Process
236
Done?
238
Yes
No
Access entry in Third Lookup Array and Process
240
Use identified OANET structure in forwarding packet
242
Return

Figure 17

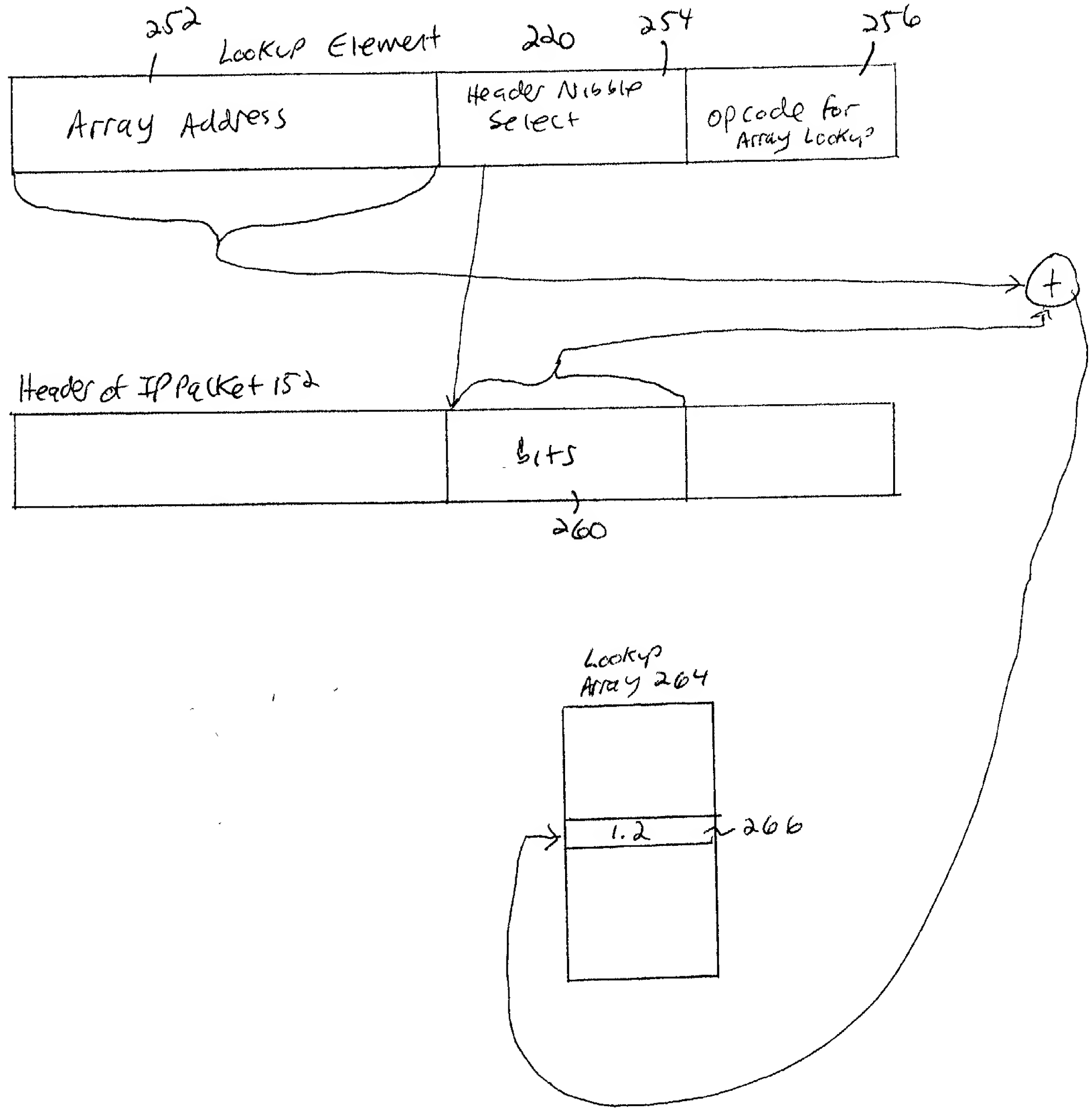

252
Lookup Element
220
254
256
Array Address
Header Nibble Select
opcode for Array Lookup
Header of IP Packet 152
bits
260
+
Lookup Array 264
1.2
266

Figure 18

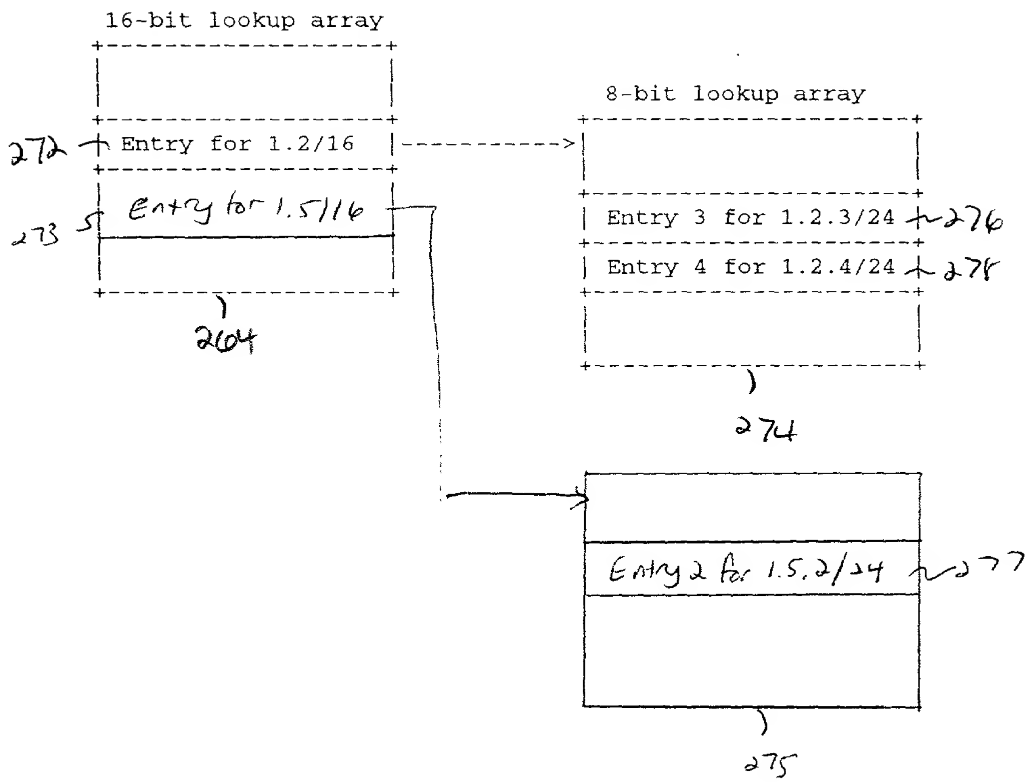

16-bit lookup array
272
Entry for 1.2/16
273
Entry for 1.5/16
264
8-bit lookup array
Entry 3 for 1.2.3/24
276
Entry 4 for 1.2.4/24
278
274
Entry 2 for 1.5.2/24
277
275

Figure 19A

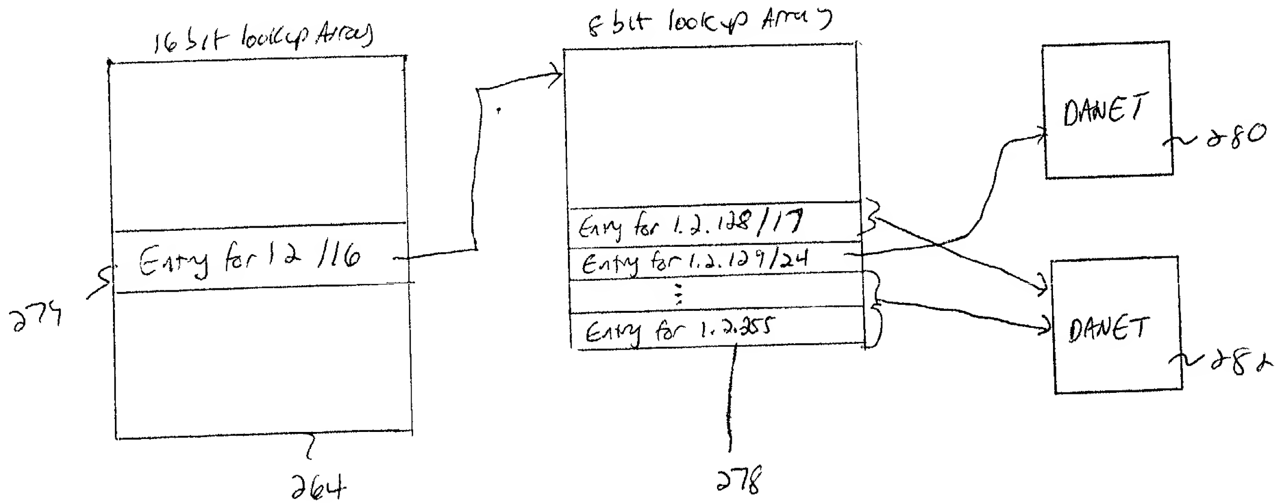

16 bit lookup Array
Entry for 1.2 /16
274
264
8 bit lookup Array
Entry for 1.2.128/17
Entry for 1.2.129/24
Entry for 1.2.255
278
DANET
280
DANET
282

Figure 19B

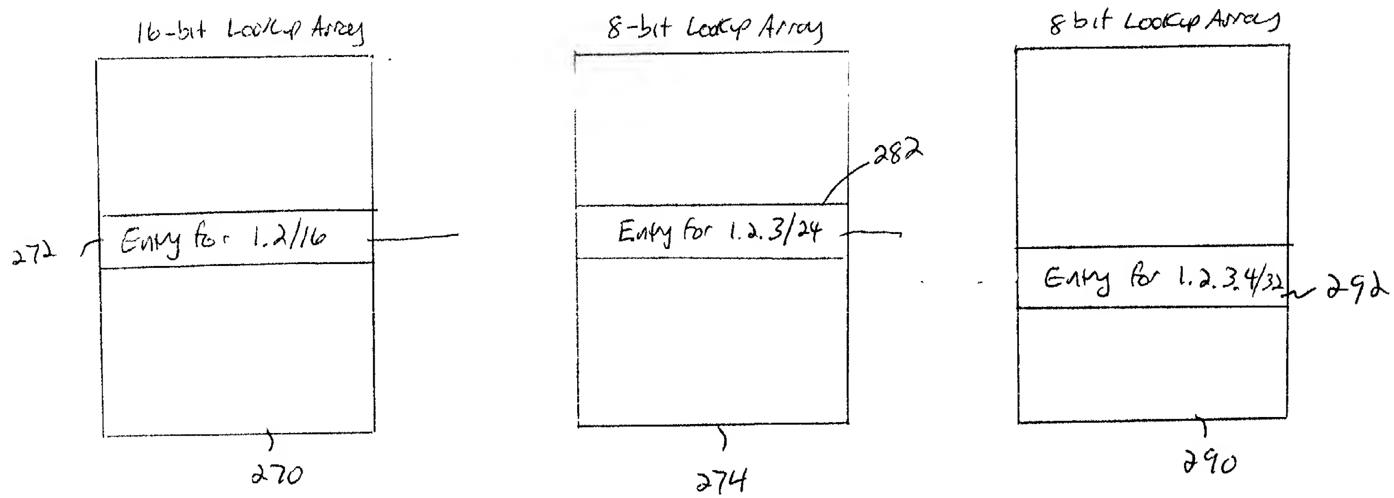

16-bit Lookup Array
Entry for 1.2/16
272
270
8-bit Lookup Array
Entry for 1.2.3/24
282
274
8 bit Lookup Array
Entry for 1.2.3.4/32
292
290

Figure 19C

Lookup Element 300

Array Address	Header Nibble Select	Opcode

Figure 20

Customer Number: 000959

Attorney's
Docket
Number AGM-002

Declaration, Petition and Power of Attorney for Patent Application

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

NETWORK PACKET FORWARDING LOOKUP WITH A REDUCED NUMBER OF MEMORY ACCESSES

the specification of which

(check one)

X is attached hereto.

_ was filed on ______________________________ as

Application Serial No. ______________________________

and was amended on ______________________________.
(if applicable)

I do not know and do not believe that the subject matter of this application was known or used by others in the United States or patented or described in a printed publication in any country before my invention thereof, or patented or described in a printed publication in any country or in public use or on sale in the United States more than one year prior to the date of this application, or first patented or caused to be patented or made the subject of an inventor's certificate by me or my legal representatives or assigns in a country foreign to the United States prior to the date of this application on an application filed more than twelve months (six months if this application is for a design) before the filing of this application; and I acknowledge my duty to disclose information of which I am aware which is material to the examination of this application, that no application for patent or inventor's certificate on the subject matter of this application has been filed by me or my representatives or assigns in any country foreign to the United States, except those identified below, and that I have reviewed and understand the contents of the specification, including the claims as amended by any amendment referred to herein.

I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

CLAIM OF BENEFIT OF EARLIER FOREIGN APPLICATION(S)

I hereby claim priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below, and have also identified below any foreign application(s) for patent or inventor's certificate filed by me on the same subject matter having a filing date before that of the application(s) from which priority is claimed.

Check one:

X no such applications have been filed.

_ such applications have been filed as follows

EARLIEST FOREIGN APPLICATION(S), IF ANY, FILED WITHIN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO THIS U.S. APPLICATION

Country	Application Number	Date of Filing (month,day,year)	Priority Claimed Under 35 USC 119
			_ Yes No _
			_ Yes No _
			_ Yes No _
			_ Yes No _
			_ Yes No _

ALL FOREIGN APPLICATION(S), IF ANY FILED MORE THAN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO THIS U.S. APPLICATION

CLAIM FOR BENEFIT OF U.S. PROVISIONAL APPLICATION(S)

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below.

60/090,028	June 19, 1998
(Application Serial No.)	(Filing Date)
(Application Serial No.)	(Filing Date)

CLAIM FOR BENEFIT OF EARLIER U.S./PCT APPLICATION(S)

I hereby claim the benefit under Title 35, United States Code, §120 of any earlier United States application(s) or PCT international application(s) designating the United States listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the earlier application(s) in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date(s) of the earlier application(s) and the national or PCT international filing date of this application. As to subject matter of this application which is common to my earlier application(s), if any, described below, I do not know and do not believe that the same was known or used by others in the United States or patented or described in a printed publication in any country before my invention thereof, or patented or described in a printed publication in any country or in public use or on sale in the United States more than one year prior to the date(s) of said earlier application(s), or first patented or caused to be patented or made the subject of an inventor's certificate by me or my legal representatives or assigns in a country foreign to the United States prior to the date(s) of said earlier application(s) on an application filed more than twelve months (six months if this application is for a design) before the filing of said earlier application(s); and I acknowledge that no application for patent or inventor's certificate on said subject matter has been filed by me or my representatives or assigns in any country foreign to the United States except those identified herein.

(Application Serial No.)	(Filing Date)	(Status) (patented,pending,aband.)
(Application Serial No.)	(Filing Date)	(Status) (patented,pending,aband.)

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorneys and/or agents to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

W. Hugo Liepmann	Reg. No. 20,407	Lawrence E. Monks	Reg. No. 34,224
James E. Cockfield	Reg. No. 19,162	David A. Lane, Jr.	Reg. No. 39,261
Thomas V. Smurzynski	Reg. No. 24,798	Catherine J. Kara	Reg. No. 41,106
Ralph A. Loren	Reg. No. 29,325	Linda M. Chinn	Reg. No. 31,240
Giulio A. DeConti, Jr.	Reg. No. 31,503	Faustino A. Lichauco	Reg. No. 41,942
Ann Lamport Hammitte	Reg. No. 34,858	Jeanne M. DiGiorgio	Reg. No. 41,710
Elizabeth A. Hanley	Reg. No. 33,505	Megan E. Williams	Reg. No. 43,270
Amy E. Mandragouras	Reg. No. 36,207	Nicholas P. Triano III	Reg. No. 36,397
John V. Bianco	Reg. No. 36,748	Peter C. Lauro	Reg. No. 32,360
Anthony A. Laurentano	Reg. No. 38,220	Reza Mollaaghababa	Reg. No. P43,810
Jane E. Remillard	Reg. No. 38,872	Timothy J. Douros	Reg. No. 41,716
Jeremiah Lynch	Reg. No. 17,425	John L. Welch	Reg. No. 28,129
Kevin J. Canning	Reg. No. 35,470		

Send Correspondence to Kevin J. Canning, Esq. at **Customer Number: 000959** whose address is:

Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109

Direct Telephone Calls to: (name and telephone number)

Kevin J. Canning, Esq., (617) 227-7400

Wherefore I petition that letters patent be granted to me for the invention or discovery described and claimed in the attached specification and claims, and hereby subscribe my name to said specification and claims and to the foregoing declaration, power of attorney, and this petition.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor Frank Kastenholz	
Inventor's signature	Date 22 JAN 99
Residence 255 North Road, #118, Chelmsford, MA 01824	
Citizenship United States	
Post Office Address (if different) Same as Above	